PadhAI: Representation Power of Functions

One Fourth Labs

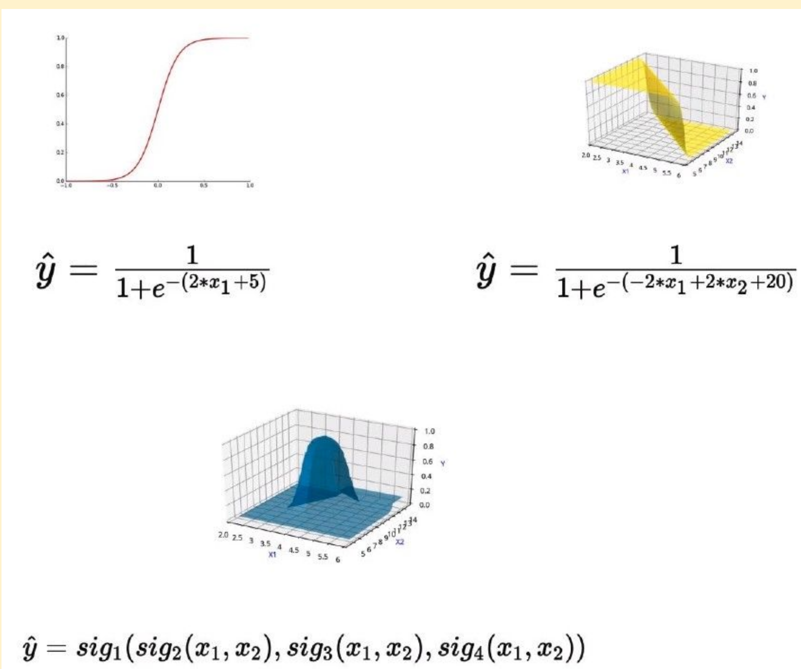
Why do we need complex functions

The need for complex functions

1. Here's a quick recap on what we've covered so far

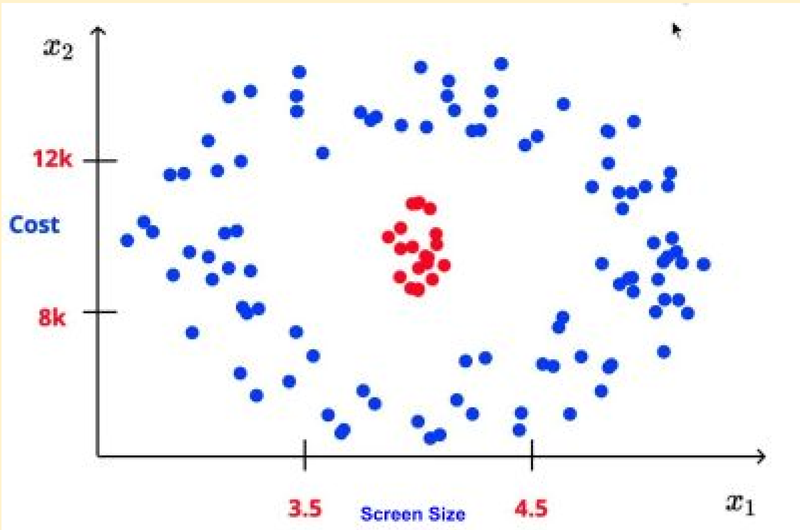
	Data	Task	Model	Loss	Learning	Evaluation
MP Neuron	{0,1}	Binary Classification	$g(x) = \sum_{i=1}^{n} x_i$ y = 1 if g(x) >= b y = 0 otherwise	Loss = $\sum_i (y_i != \hat{y}_i)$	Brute Force Search	Accuracy
Perceptron	Real Inputs	Binary Classification	y = 1 if $\sum_{i=1}^{n} w_i x_i >= b$ y = 0 otherwise	Loss = $\sum_i (y_i - \hat{y}_i)^2$	Perceptron Learning Algorithm	Accuracy
Sigmoid	Real Inputs	Classification /Regression	$y = \frac{1}{1 + e^{-(w^T x + b)}}$	Loss = $\sum_i (y_i - \hat{y}_i)^2$ Or Loss = $-[(1-y)log(1-\hat{y}) + ylog(\hat{y})]$	Gradient Descent	Accuracy/RMSE

2. We must remember that none of the above 3 models can handle non-linearly separable data
3. Here's another recap on Continuous Functions

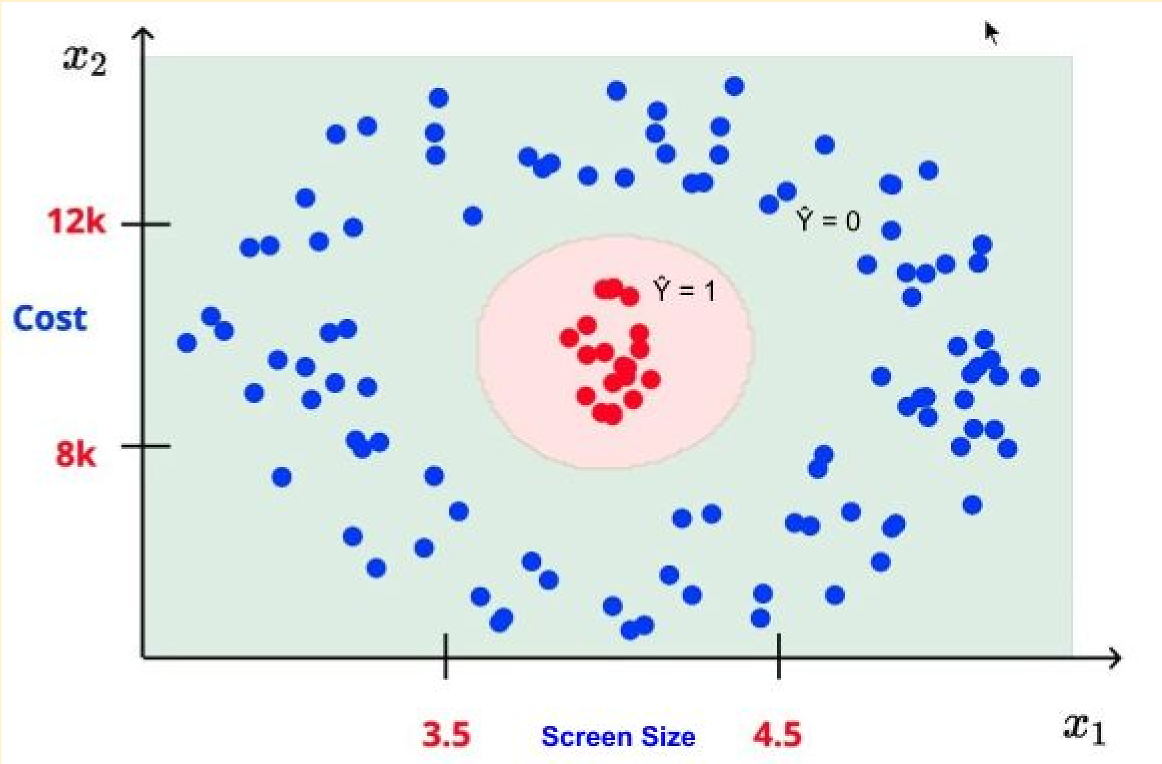


4. We care about continuous functions because our learning algorithm (Gradient Descent) requires that the input functions be differentiable (i.e. Continuous)
5. Let's take a look at a real world example of how complex functions are relevant to our situation

6. Consider the following example of where we're trying to predict like/dislike for a non-linearly separable dataset of mobile phones.

7. 

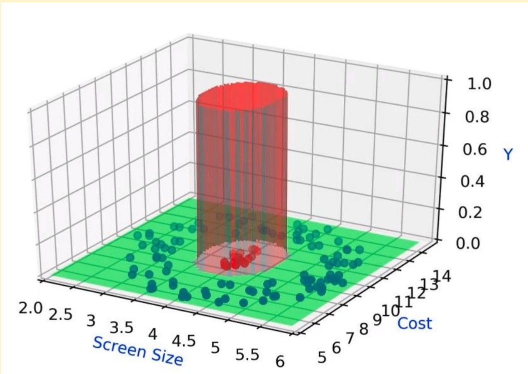
8. Here, our desirable set of phones lies in the centre of a circle of non-desirable phones, based on the values of the variable Cost and Screen Size.
9. Ideally, we would need a decision boundary like so

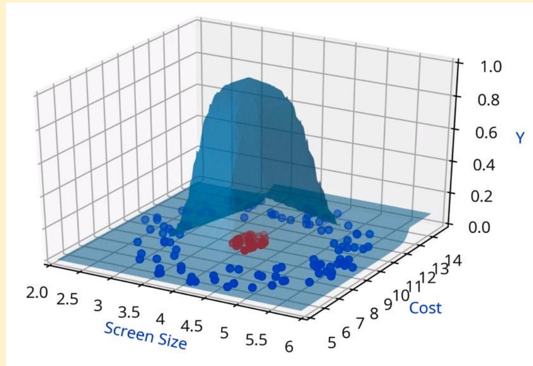


10. However, none of the functions we have seen so far will be able to plot such a decision boundary (ie boundary that separates the two classes = 0 and ŷ = 1)

11. Let's take a 3D plot of the two variables with the output values mapped along the z-axis

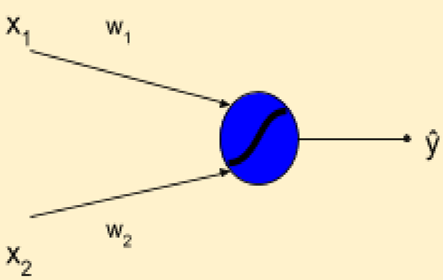
Discrete (abrupt) **Continuous (smooth)**

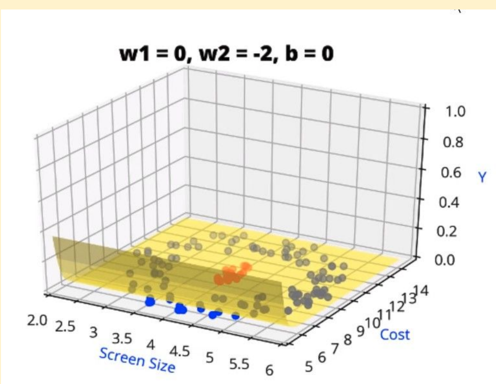


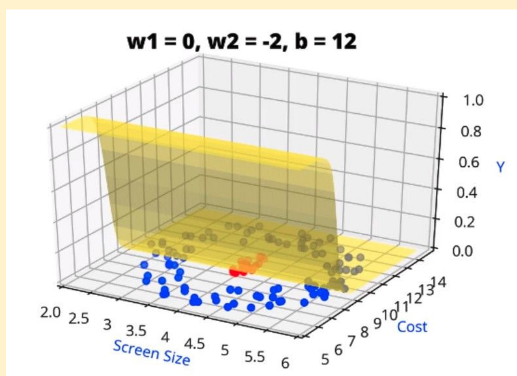


12. Here, the Continuous function has a smooth distribution, and the Y value gradually increases as we converge to the centre, becoming 1 at the region around the red dots
13. However, such an output is not possible with the sigmoid functions, regardless of how we manipulate the values of w and b

Sigmoid decision boundary, can range from s-shape to flat, based on w and b values







14. We can see that the sigmoid function is unsuitable for modelling complex decision boundaries.
15. Such complex relations are actually seen quite frequently in real world examples